AP Statistics Practice Test (page 468)

T7.1 c. The statistic is a measure of the sample and the parameter is a measure of the population.

T7.2 c. Sample size has no effect on the bias of an estimate, but larger samples will reduce the variability of an estimate.

T7.3 c. To use the Normal approximation, both np and $n(1-p)$ must be at least 10. In all options other than c, this condition is not met.

T7.4 a. The central limit theorem applies when the sample size is large ($n \geq 30$). The CLT is not needed when the original distribution is Normal; the distribution of the sample mean is always Normal in that case.

T7.5 b. Because a sample of 3% of the undergraduates from Ohio State University consists of approximately 1800 students whereas a sample of 3% of the undergraduates from Johns Hopkins consists of just 60 students, the estimate from Ohio State University will have less sampling variability.

T7.6 b. Because the standard deviation of the sampling distribution of the sample mean is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, doubling the sample size gives $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{2n}} = \frac{1}{\sqrt{2}}\frac{\sigma}{\sqrt{n}}$.

T7.7 b. The sampling distribution would be only approximately Normal with mean equal to the population proportion (0.55 in this case) and standard deviation equal to $\sqrt{\frac{(0.55)(0.45)}{250}} = 0.03$.

T7.8 e. The sampling distribution has information about how the sample mean varies from sample to sample, not what any sample itself looks like.

T7.9 c. $\mu_{\bar{x}} = \mu = 16.05$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.1}{\sqrt{4}} = 0.05$. About 95% of the sample means should be within two standard deviations of the mean, which is $16.05 \pm 2(0.05) = 15.95$ to 16.15.

T7.10 e. The distribution of the average amount of pay will not be Normal because there are only three possible outcomes, $\bar{x}$ = 40, 60, or 80. $\bar{x}$ = 40 will occur 25% of the time, as will $\bar{x}$ = 80. $\bar{x}$ = 60 will happen 50% of the time.

T7.11 Sample statistic A provides the best estimate of the parameter. Both statistics A and B appear to be unbiased, while statistic C appears to be biased because the center of its sampling distribution is smaller than the value of the parameter. In addition, statistic A has lower variability than statistic B. In this situation, we want low bias and low variability, so statistic A is the best choice.

T7.12 (a) The probability that a single household pays more than \$39 cannot be calculated, because we do not know the shape of the population distribution of monthly fees.
(b) The mean of the sampling distribution of the sample mean is equal to the mean of the population distribution. Therefore $\mu_{\bar{x}} = \mu = \$38$. Also we know that $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{10}{\sqrt{500}} = \$0.4472$ because the sample of size 500 is less than 10% of all households with Internet access.
(c) Because the sample size is large ($n = 500 \geq 30$), the distribution of $\bar{x}$ will be approximately Normal.
(d) **Step 1: State the distribution and values of interest.** We want to find $P(\bar{x} > 39)$ using the $N(38, 0.4472)$ distribution. **Step 2: Perform calculations. Show your work.** The standardized score for the boundary value is $z = \dfrac{39-38}{0.4472} = 2.24$. The desired probability is $P(Z > 2.24) = 1 - 0.9875 = 0.0125$. *Using technology*: normalcdf(lower: 39, upper: 1000, μ : 38, σ : 0.4472) = 0.0127. **Step 3: Answer the question.** There is a 0.0127 probability that the mean monthly fee exceeds \$39.

T7.13 **Step 1: State the distribution and values of interest.** We have an SRS of size 300 drawn from a population in which the proportion who live in households with incomes under the poverty line is $p = 0.22$. Thus, $\mu_{\hat{p}} = p = 0.22$. Because 300 is less than 10% of children under the age of six, $\sigma_{\hat{p}} = \sqrt{\dfrac{0.22(0.78)}{300}} = 0.0239$. Because $np = 300(0.22) = 66$ and $n(1-p) = 300(0.78) = 234$ are both at least 10, the sampling distribution of $\hat{p}$ can be approximated by a Normal distribution. We want to find $P(\hat{p} > 0.20)$ using the $N(0.22, 0.0239)$ distribution. **Step 2: Perform calculations. Show your work.** The standardized score for the boundary value is $z = \dfrac{0.20-0.22}{0.0239} = -0.84$. The desired probability is $P(Z > -0.84) = 1 - 0.2005 = 0.7995$ *Using technology*: normalcdf (lower: 0.20, upper: 1000, μ : 0.22, σ : 0.0239) = 0.7987. **Step 3: Answer the question.** There is a 0.7987 probability that more than 20% of the sample are from poverty-level households.